Hello harlie

Enabling speech monitoring through chat-bot conversations

Ireland, David; Atay, Christina; Liddle, Jacki; Bradford, Dana; Lee, Helen; Rushin, Olivia; Mullins, Thomas; Angus, Dan; Wiles, Janet; McBride, Simon; Vogel, Adam

Published in:
Digital Health Innovation for Consumers, Clinicians, Connectivity and Community - Selected Papers from the 24th Australian National Health Informatics Conference, HIC 2016

DOI:
10.3233/978-1-61499-666-8-55

Digital Health Innovation for Consumers, Clinicians, Connectivity and Community
A. Georgiou et al. (Eds.)

doi:10.3233/978-1-61499-666-8-55

Hello Harlie: Enabling Speech Monitoring Through Chat-Bot Conversations

David IRELAND [a,1], Christina ATAY [b] Jacki LIDDLE [b], Dana BRADFORD [a], Helen LEE [c], Olivia RUSHIN [c], Thomas MULLINS [c], Dan ANGUS [a], Janet WILES [c], Simon MCBRIDE [a], Adam VOGEL [d]

[a] *The Australian E-Health Research Centre, CSIRO, Australia*
[b] *Asia Pacific Center of Neuromodulation, University of Queensland, Australia*
[c] *School of ITEE, The University of Queensland, Australia*
[d] *Centre for Neuroscience of Speech, The University of Melbourne, Australia*

Abstract. People with neurological conditions such as Parkinson's disease and dementia are known to have difficulties in language and communication. This paper presents initial testing of an artificial conversational agent, called *Harlie*. Harlie runs on a smartphone and is able to converse with the user on a variety of topics. A description of the application and a sample dialog are provided to illustrate the various roles chat-bots can play in the management of neurological conditions. Harlie can be used for measuring voice and communication outcomes during the daily life of the user, and for gaining information about challenges encountered. Moreover, it is anticipated that she may also have an educational and support role.

Keywords. neurodegenerative diseases, artificial intelligence, natural language processing, speech, voice, remote monitoring.

Introduction

Many people living with neurological conditions like autism, Parkinson's disease or dementia have difficulties communicating or conversing. Taking Parkinson's disease as an example, voice changes are experienced by 80-90% of people [1], who may also subsequently experience changes to their thinking, use of language and mood [2]. Early research focused largely on the various phonation impairments, phoneme articulation and acoustic timing [3,4]. More recent efforts, however, have been placed on examining the consequences of difficulties with communication. Studies have shown people with Parkinson's disease (PwPD) have problems with conversation initiation, turn-taking, topic management, word-retrieval, and memory [5]. A notable article by Miller *et al.* in [1] examined the impact of changes in communication by conducting in-depth interviews and found emergent themes of frustration due to losing track of thoughts mid-sentence and indignity and social withdrawal from being excluded from the conversations.

With the advent of ubiquitous smart-phones, autonomous remote monitoring applications that log and analyse multi-domain data such as accelerometery, voice and com-

[1] Corresponding Author: David Ireland; E-mail: d.ireland@csiro.au

munity movement are being increasingly reported [6]. None however, has yet considered the monitoring and logging of conversation. This is perhaps due to the complex ethical and privacy issues surrounding autonomous audio recording.

In earlier work [7], we proposed the concept of a conversation agent, or chat-bot, tailored for remote monitoring of audio and conversation dialogues. Here we detail the realisation of this proposal in the form of a smartphone application referred to as *HARLIE* (Human And Robot Language Interaction Experiment). At the moment Harlie runs on the Android operating system and its purpose is to chat about various topics whilst analysing the user's voice and language articulation. At the moment, Harlie is available freely on the Android playstore [8] as part of a research project targeting the general population, which has approval from an institutional ethics committee and commenced in December 2015.

Chat-bots have been reported in the literature for health related applications. Examples include health behavior change for obesity and diabetes [9], disease self-management [10]; and health education for adolescents on topics related to sex, drugs and alcohol [11]. Our work is significantly different in two ways. First, technology that can monitor progress and the impact of difficulties with communication could help the person, as well as health professionals and researchers, understand the impact of health conditions on communication. Second, technology can provide people with a way to privately practise their communication without feeling frustrated or judged while receiving encouragement and individualised feedback.

1. Harlie the Chatbot

1.1. User Interaction

Upon installing Harlie, the user is presented with the consent form and subsequently a dialog screen asking for demographic information as shown in Figure 1a. Harlie has been programmed to randomly call the user once a day, between 8AM - 8PM (Figure 1b), however users can also initiate a conversation at any time. To talk to Harlie the user is required to hold down the bottom green button whilst speaking - this ensures no unintended audio is captured (Figure 1c). Harlie processes the input and responds via speaking and displaying the text on the screen. The user can end the conversation at any time by pressing the *Exit* button. Harlie can also end the conversation when she has captured a pre-set minimum of audio. If Harlie decides to end the conversation, it waits for an appropriate time to end the conversation, usually at the end of the topic of discussion. After exiting the user is asked whether the audio and dialog can be remotely logged (Figure 1d) and subsequently whether Harlie's responses were consistent (Figure 1e). If the users wishes to withdraw from the study, a dialog screen is accessible that allows the user to withdraw from the study and have their data deleted from the server; alternatively the user can withdraw from the study and allow the data to remain (Figure 1f).

1.2. Harlie's Brain

Early reported chat-bots solely relied on typed text as the input stimuli, however with modern smartphone technology, this need no longer be the case. Android and iPhone

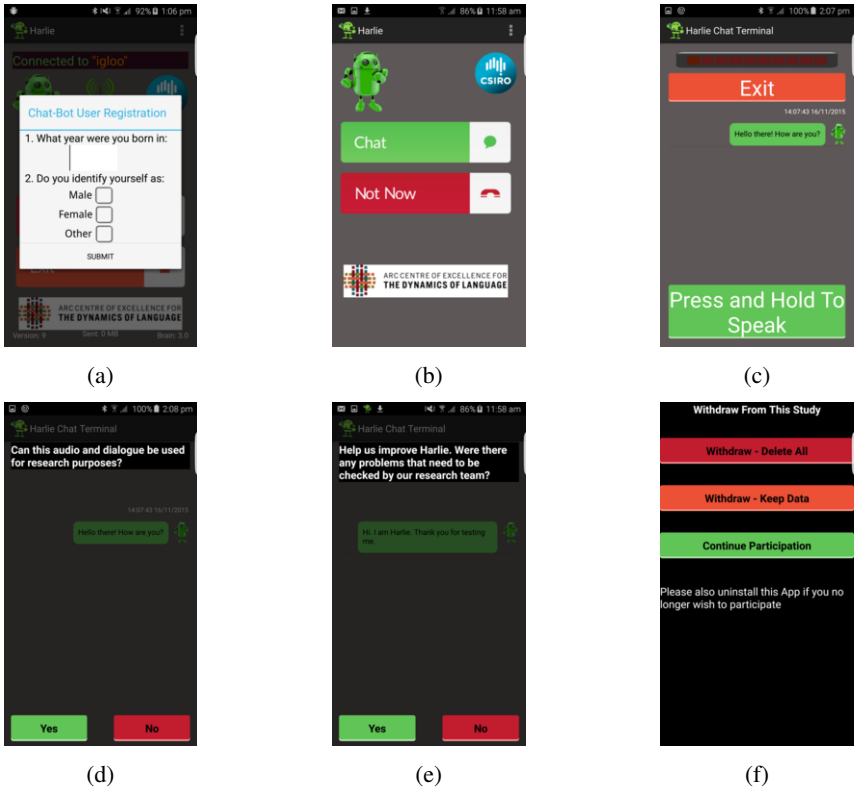

Figure 1. Screenshots of the Harlie application. (a) The registration screen; (b) the calling screen; (c) the chatting terminal (d) the screen asking for user permission to remote log the conversation; (d) the screen giving the user options for withdrawing from the study.

based smartphones have both speech-to-text and text-to-speech tools capable of converting spoken acoustic signals into digital text and converting digital text into a digital-synthetic, acoustic voice. Harlie uses Google's speech-to-text and text-to-speech application programming index (API) to perform these activities. Although this requires sending the acoustic audio to an off-shore server, a random voice modulation is applied to the signal prior to being sent. This ensures the user cannot be identified by their speech pattern if intercepted.

How the text input is processed to construct a meaningful response has been the main research question of chat-bot researchers for some time. The most prominent technique for the last decade has been via case-based reasoning and textual pattern matching algorithms in particular the use of a standardised computer language referred to as artificial intelligence mark-up language (AIML). Harlie uses AIML to converse with a user in a way that is both non-deterministic and meaningful. Her so-called digital *brain* is made of collections of AIML files for various topics, situations and speech tasks.

An exhaustive list of AIML features is beyond the scope of this paper and the reader is referred to [12] for more information. It is worth noting that to add new topics, the prospective author does not need to have any programming skills. Indeed, the majority of Harlie's brain was written by speech pathologists and language psychologists who

have developed modules for numerous topics including for example, family, pets, work, travelling and music.

1.3. Speech Tasks

During the chat, Harlie actively analyses aspects of the health of users' voice and communication. This includes how well vowels are articulated, vocabulary range, and duration of mid-sentence pauses. People who might be working on improving their voice or communication due to difficulties related to a neurological condition, like Parkinsons disease or stroke, may need to practise and get feedback on a daily basis.

2. Focus Group Study

2.1. Feedback

In addition to use by the general public, several Harlie focus group studies were carried out in Brisbane. Existing community groups, which involved older people, were invited to host a visit from the research team. Participants were provided background information regarding Harlie and asked to sign consent forms prior to participation. Participants were given time to speak with Harlie and then subsequently asked to share their experiences. They could also opt out of sharing their conversations with the research team.

Individual and focus group feedback was obtained from 33 participants (17 females and 16 males) aged 27 to 87 (mean age: 66.5 years). Participants ranged in their experience with technology with 70% owning a smartphone and 18% describing themselves as competent users of Android phones. Some of the community participants also described themselves as competent users of iphones (36%) and tablet computers (48%), laptops (42%) and desktop computers (58%).

Feedback included initial impressions, difficulties encountered and practical applications. Overall the impression from the first use of Harlie was positive, with participants also identifying technical problems (speed of processing), problematic conversational responses and potential future uses for Harlie. Technical and conversational issues were addressed in improving those aspects of the app. A frequent suggestion was providing company for people in residential aged care facilities, as described by a participant:

"In old homes, people are very lonely, there is no people to talk around. If you can just install that program, people would love to talk about their family, their children, their grandchildren ..."

2.2. Example Chat

A dialog excerpt between Harlie and a member of one of the community groups, who coincidentally has autism, is given in Figure 2. This excerpt shows the user discussing what woodwork is done at the particular community centre the user attends. A coloured rectangle is placed next to each utterance spoken by the user and Harlie. The length of the rectangle reflects the time duration of the utterance. The placement of the rectangles corresponds to when the utterance occurs in the timeline of the conversation. This allows

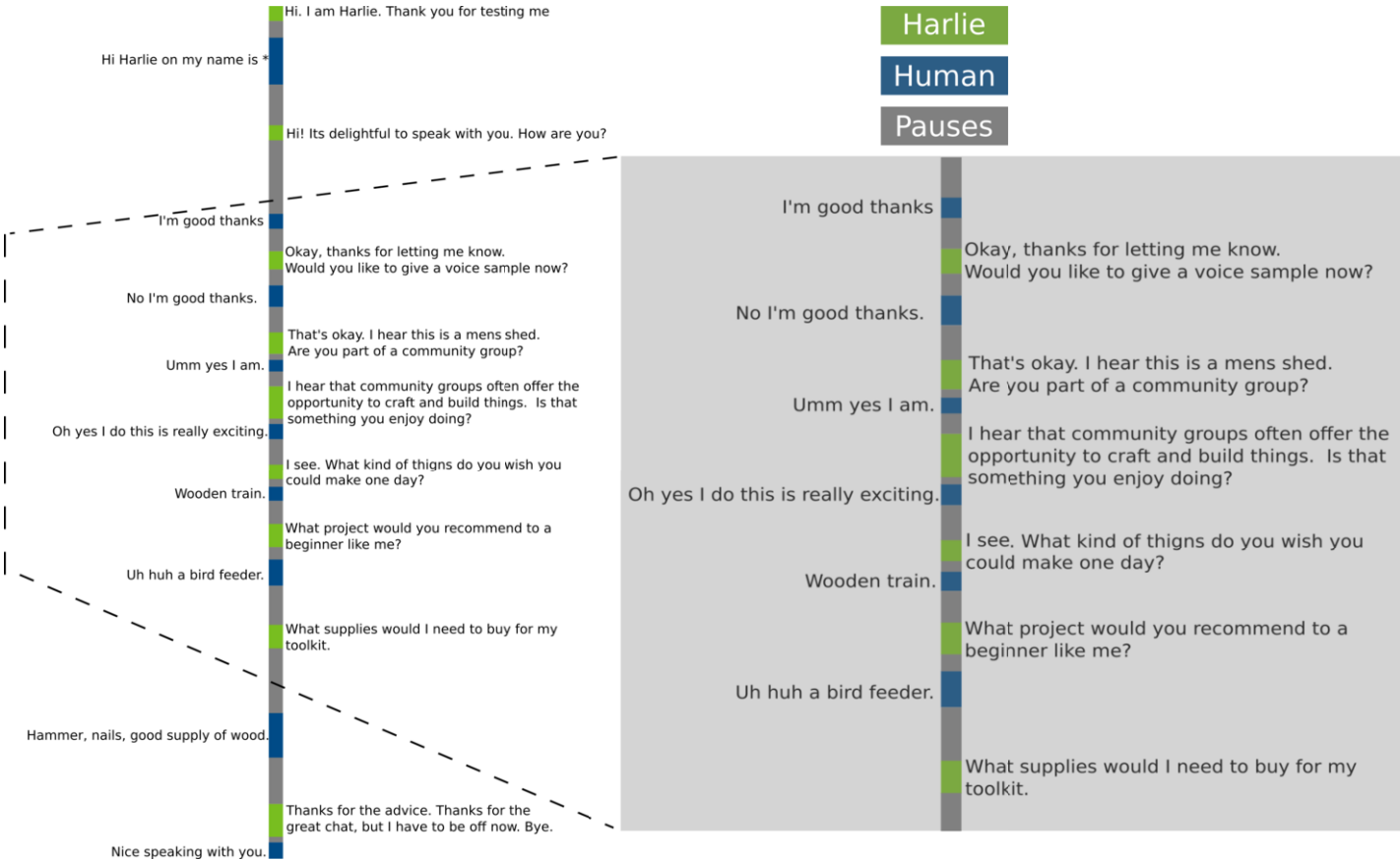

Figure 2.: A sample of dialog between Harlie and a person with autism. Green rectangles designate dialog generated from Harlie; blue rectangles designate responses from the user; grey rectangles indicate pauses in the timeline.

the visualisation of gaps or silences that occur throughout the conversation. As Harlie relies on the Internet to process the speech to text, random latency does occur ranging from milliseconds to a few seconds. Silences from the user could arise from a multitude of issues such as being distracted but may also be due to clinical conditions including speech impediments and memory impairment.

Figure 2 shows that gaps are apparent in the first several utterances. This could be attributed to the user feeling unsure or uncomfortable with using the technology for the first time. When the topic was changed to focus on the community group, a much more rapid interchange took place. The participant was able to convey his desires not to give a speech sample and describe, albeit briefly his interests in woodwork.

3. Discussion

This paper has provided an overview of the workings of smartphone application called Harlie. Harlie was developed to converse with the user on various topics of general conversation and to facilitate assessments. As more people speak to her, Harlie's repertoire of topics will expand. This work begins to develop and employ a conversation agent that mimics human conversation using contemporary technology. Given the practical benefits of the chat-bot, it could easily play a role in the next generation of speech and communication therapy for people living with many neurological and other conditions.

References

[1] N. Miller, E. Noble, D. Jones, D. Burn, *Life with communication changes in Parkinson's disease*, Age and Ageing 35(3):235-239 Epub 2006

[2] D. Robertson, *Maintaining the art of conversation in Parkinsons disease*. Age Ageing 2006 35: 211

[3] S. Sapir, A.A. Pawlas, L.O. Ramig *et al.*, *Voice and speech abnormalities in Parkinson disease: Relation to severity of motor impairment, duration of disease, medication, depression, gender and age*, J Med Speech-Lang Pathol 2001;9:213-26

[4] K. Forrest, G. Weismer, G.S. Turner, *Kinematic, acoustic, and perceptual analyses of connected speech produced by parkinsonian and normal geriatric adults.*, J Acoust Soc Am. 1989 Jun;85(6):2608-22.

[5] C. Saldert, U. Ferm, S. Bloch. *Semantic trouble sources and their repair in conversations affected by Parkinsons disease*, International Journal of Language & Communication Disorders 2014;49(6):710-721. doi:10.1111/1460-6984.12105.

[6] J. Liddle, D. Ireland, S.J. McBride, S.G. Brauer, L.M. Hall, H. Ding, M. Karunanithi, P.W. Hodges, D. Theodoros, P.A. Silburn, H.J. Chenery, *Measuring the lifespace of people with Parkinson's disease using smartphones: proof of principle*, JMIR Mhealth Uhealth. 2014 Mar 12;2(1):e13

[7] D. Ireland, J. Liddle, S. McBride, H. Ding, and C. Knuepffer, *Chat-Bots for People with Parkinson's Disease: Science Fiction or Reality?*, Studies in Health Technology and Informatics, Volume 214: Driving Reform: Digital Health is Everyones Business

[8] Harlie The Chatbot. [online], `https://play.google.com/store/apps/details?id=org.harlie.thechatbot&hl=en`, Last Accessed: 05/02/2016

[9] A. Watson, T. Bickmore, A. Cange, A. Kulshreshtha, J. Kvedar *An internet-based virtual coach to promote to physical activity adherence in overweight adults: randomized controlled trial*, J. Med Internet Res. 2012 Jan-Feb 14(1)

[10] T. Bickmore, A. Gruber, R. Picard, *Establishing the computer-patient working alliance in automated health behavior change interventions*, Patient Education and Counselling 2005 Oct;59(1):21-30.

[11] R. Crutzen, G.Y. Peters, S. D. Portugal, E. M. Fisser, J.J Grolleman, *An Artificially Intelligent Chat Agent That Answers Adolescents' Questions Related to Sex, Drugs, and Alcohol: An Exploratory Study*, Journal of Adolescent Health , Volume 48 , Issue 5 , 514 - 519

[12] R. Wallace, *The Elements of AIML Style*, ALICE A.I. Foundation, 2001